

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 10-171497

(43)Date of publication of application : 26.06.1998

(51)Int.Cl. G10L 9/00

G06F 17/10

H03H 17/00

H04B 15/00

(21)Application number : 08-332182

(71)Applicant : OKI ELECTRIC IND CO LTD

(22)Date of filing : 12.12.1996

(72)Inventor : TAKADA SHINSUKE

ARIYAMA YOSHIHIRO

(54) BACKGROUND NOISE REMOVER DEVICE

(57) Abstract:

PROBLEM TO BE SOLVED: To provide a background noise remover device which, when receiving input signals where background noises and speech signals are mixed, can remove the background noise without a damage on sound quality.

SOLUTION: While input signals do not include background noises, a background noise estimation unit 7 produces and holds therein an estimated background noise $N2_{f,k}$. When input signals include background noises, an adder 5 calculates $E1_{f,k}$ with an equation: $E1_{f,k} = X_{f,k} - \alpha_c N2_{f,k}$, where $N2_{f,k}$ is frequency component of the input signals. In order to suitably process a channel having a smaller S/N than other channels, a unit 12 produces a re-updated estimated background noise $N3_{mf,k}$ and an adder 13 carries out subtraction as the following: $E2_{mf,k} = X_{mf,k} - \alpha_c N2_{mf,k} - \alpha_k N3_{mf,k}$, where α_k is a predefined coefficient.

(19)日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平10-171497

(43)公開日 平成10年(1998) 6月26日

(51)Int.Cl.⁶

識別記号

F I

G 1 0 L 9/00

G 1 0 L 9/00

F

G 0 6 F 17/10

H 0 3 H 17/00

6 0 1 M

H 0 3 H 17/00

6 0 1

H 0 4 B 15/00

H 0 4 B 15/00

G 0 6 F 15/31

D

審査請求 未請求 請求項の数 5 O L (全 12 頁)

(21)出願番号

特願平8-332182

(22)出願日

平成8年(1996)12月12日

(71)出願人 000000295

沖電気工業株式会社

東京都港区虎ノ門1丁目7番12号

(72)発明者 高田 真資

東京都港区虎ノ門1丁目7番12号 沖電気
工業株式会社内

(72)発明者 有山 義博

東京都港区虎ノ門1丁目7番12号 沖電気
工業株式会社内

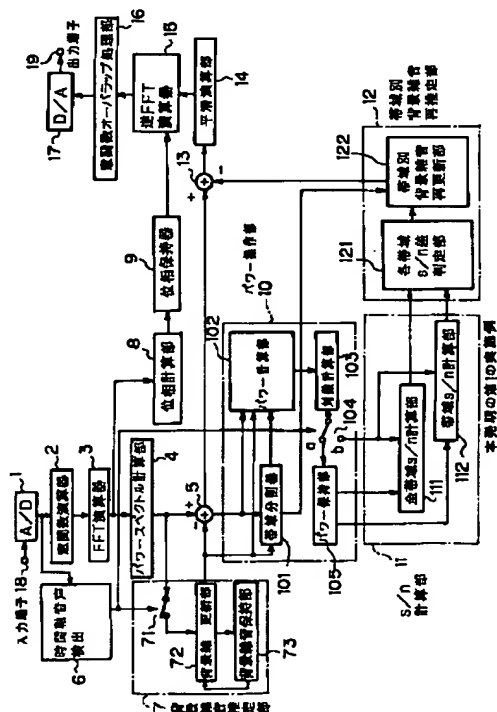
(74)代理人 弁理士 香取 孝雄

(54)【発明の名称】 背景雑音除去装置

(57)【要約】

【課題】 背景雑音と音声信号が混在する入力信号から背景雑音を音質を損なうことなく除去し得る背景雑音除去装置を提供する。

【解決手段】 A/D変換器1～パワースペクトル計算部4で入力信号を順次フレーム単位の周波数成分 $X_{f,k}$ に変換する。時間軸音声検出器6が音声信号を検出しない期間、背景雑音推定部7は $X_{f,k}$ に基づいて推定背景雑音 $N_{2f,k}$ を生成して保持する。時間軸音声検出器6が音声信号を検出する期間、加算器5は $E1_{f,k} = X_{f,k} - \alpha_c N_{2f,k}$ を計算する。パワー操作部10とS/N計算部11は、 $E1_{f,k}$ と $N_{2f,k}$ に基づいて全周波数帯域の S/N_a と全周波数帯域を複数に分割した小領域毎の S/N_m とを計算する。帯域別背景雑音再推定部12は、 S/N_m と S/N_a の差が所定値以下の各小領域について再更新推定背景雑音 $N_{3mf,k}$ を生成する。加算器13は $E1_{f,k}$ から所定量の $N_{3mf,k}$ を減算し、平滑演算部14～D/A変換器17等は加算器13の出力をフレーム毎の周波数成分から時間軸信号に変換して出力する。



【特許請求の範囲】

【請求項 1】 入力信号として、背景雑音が入力され次いで該背景雑音と音声信号が混在する形で入力される背景雑音除去装置において、該装置は、前記入力信号を時間軸信号から順次フレーム単位で周波数成分に変換する信号変換手段と、前記入力信号に含まれる音声信号を検出する音声検出手段と、

該音声検出手段で音声信号が検出されないノイズ期間は、前記信号変換手段で変換される周波数成分と 1 フレーム前に生成した推定背景雑音との平均をとることにより推定背景雑音を生成して保持する背景雑音推定手段と、

前記音声検出手段で音声信号が検出される音声期間は、前記信号変換手段で変換される周波数成分から前記背景雑音推定手段で保持される推定背景雑音を減算する第 1 の加算手段と、

該第 1 の加算手段の減算で得られる周波数成分を信号とし前記背景雑音推定手段で保持される推定背景雑音を雑音として、全周波数帯域の S/N と全周波数帯域を複数に分割した小領域毎の S/N とを計算する S/N 計算手段と、前記小領域の S/N と全周波数帯域の S/N の差が所定値以下の各小領域について、前記第 1 の加算手段の減算で得られる周波数成分と前記背景雑音推定手段で保持される推定背景雑音とを所定の割合で含む再更新推定背景雑音を生成する帯域別背景雑音再推定手段と、

前記第 1 の加算手段の減算で得られる周波数成分から前記再更新推定背景雑音を減算する第 2 の加算手段と、

該第 2 の加算手段の減算で得られるフレーム毎の周波数成分を時間軸信号に変換して出力する信号再生手段とを有し、

前記第 1 の加算手段で減算する推定背景雑音と前記第 2 の加算手段で減算する再更新推定背景雑音の大きさは、前記入力信号から背景雑音が除去されるように設定されることを特徴とする背景雑音除去装置。

【請求項 2】 入力信号として、背景雑音が入力され次いで該背景雑音と音声信号が混在する形で入力される背景雑音除去装置において、該装置は、前記入力信号を時間軸信号から順次フレーム単位で周波数成分に変換する信号変換手段と、前記入力信号に含まれる音声信号を検出する音声検出手段と、

該音声検出手段で音声信号が検出されないノイズ期間は、前記信号変換手段で変換される周波数成分と 1 フレーム前に生成した推定背景雑音との平均をとることにより推定背景雑音を生成して保持する背景雑音推定手段と、

前記音声検出手段で音声信号が検出される音声期間は、前記信号変換手段で変換される周波数成分から前記背景雑音推定手段で保持される推定背景雑音を減算する第 1

の加算手段と、

該第 1 の加算手段の減算で得られる周波数成分を信号とし前記背景雑音推定手段で保持される推定背景雑音を雑音として、全周波数帯域の S/N と全周波数帯域を複数に分割した小領域毎の S/N とを計算する S/N 計算手段と、前記小領域の S/N と全周波数帯域の S/N の差が所定値以下の各小領域について、前記第 1 の加算手段の減算で得られる周波数成分と前記背景雑音推定手段で保持される推定背景雑音とを所定の割合で含む雑音を生成し、該雑音に前記背景雑音推定手段で保持される推定背景雑音を加算することにより再更新推定背景雑音を生成する帯域別背景雑音再推定手段と、

前記信号変換手段の変換で得られる周波数成分から前記再更新推定背景雑音を減算する第 2 の加算手段と、

該第 2 の加算手段の減算で得られるフレーム毎の周波数成分を時間軸信号に変換して出力する信号再生手段とを有し、

前記第 2 の加算手段で減算する再更新推定背景雑音の大きさは、前記入力信号から背景雑音が除去されるように設定されることを特徴とする背景雑音除去装置。

【請求項 3】 請求項 1 または請求項 2 に記載の背景雑音除去装置において、前記音声検出手段は、前記信号変換手段の変換で得られる周波数成分に含まれる音声成分を検出することにより前記入力信号に含まれる音声信号を検出することを特徴とする背景雑音除去装置。

【請求項 4】 請求項 3 に記載の背景雑音除去装置において、前記音声検出手段は、前記信号変換手段の変換で得られる周波数成分を複数に分割した小領域毎に音声成分を検出し、音声成分が検出された小領域の数が所定値以上であるとき前記入力信号に音声信号が含まれると判定することを特徴とする背景雑音除去装置。

【請求項 5】 請求項 3 に記載の背景雑音除去装置において、前記音声検出手段は、前記信号変換手段の変換で得られる周波数成分を複数に分割した小領域毎に音声成分を検出し、全小領域数に対する音声成分が検出された小領域の数の割合が所定値以上であるとき前記入力信号に音声信号が含まれると判定することを特徴とする背景雑音除去装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、音声信号と背景雑音が混在する入力信号から背景雑音を除去する背景雑音除去装置に関するものである。

【0002】

【従来の技術】音声信号と背景雑音が混在する入力信号から背景雑音を除去する一方法として、入力信号を周波数成分に変換してあらかじめ定めた低周波数成分と高周波数成分とに分け、低周波数成分に対しては全体に 1 より小さいゲインをかけて振幅を小さくし、高周波数成分に対しては周波数減算法(spectral subtraction)を適用

3

して各スペクトル毎に背景雑音を除去する方法が知られている。

【0003】

【発明が解決しようとする課題】しかしながら、上記の背景雑音除去方法では、(a) 通常の使用状態では背景雑音の周波数特性は様々で未知であるため、あらかじめ低周波数域と高周波数域とを決定しておくことは困難である。(b) 低周波数域と高周波数域とでは別々の方式を用いて処理を行うので、低周波数域と高周波数域の分割点での整合が難しく、スペクトルが低周波数域から高周波数域まで連続した周波数特性に対しても、周波数分割点において処理方式の相違による不連続な処理点が発生し、聴感上の違和感を発生させる。(c) 低周波数域ではS/Nが改善されずにパワーが小さくなるだけであるので、特に音声成分が低周波数域に集中している場合には効果的な雑音除去が期待できず、周波数成分の全帯域で見たとときには、S/Nはむしろ劣化し、音質も劣化する。等の欠点があった。

【0004】本発明はこのような従来技術の欠点を解消し、低周波数域と高周波数域を区別することなく周波数成分の全帯域で1つの方法で背景雑音の除去を行う、聴感上の違和感が生じない、音質の優れた背景雑音除去装置を提供することを目的とする。

【0005】

【課題を解決するための手段】本発明は上述の課題を解決するために、入力信号として、背景雑音が入力され次いで該背景雑音と音声信号が混在する形で入力される背景雑音除去装置において、この装置は、入力信号を時間軸信号から順次フレーム単位で周波数成分に変換する信号変換手段と、入力信号に含まれる音声信号を検出する音声検出手段と、この音声検出手段で音声信号が検出されないノイズ期間は、信号変換手段で変換される周波数成分と1フレーム前に生成した推定背景雑音との平均をとることにより推定背景雑音を生成して保持する背景雑音推定手段と、音声検出手段で音声信号が検出される音声期間は、信号変換手段で変換される周波数成分から背景雑音推定手段で保持される推定背景雑音を減算する第1の加算手段と、この第1の加算手段の減算で得られる周波数成分を信号とし、背景雑音推定手段で保持される推定背景雑音を雑音として、全周波数帯域のS/Nと全周波数帯域を複数の分割した小領域毎のS/Nとを計算するS/N計算手段と、小領域のS/Nと全周波数帯域のS/Nの差が所定値以下の各小領域について、第1の加算手段の減算で得られる周波数成分と背景雑音推定手段で保持される推定背景雑音とを所定の割合で含む再更新推定背景雑音を生成する帯域別背景雑音再推定手段と、第1の加算手段の減算で得られる周波数成分から再更新推定背景雑音を減算する第2の加算手段と、この第2の加算手段の減算で得られるフレーム毎の周波数成分を時間軸信号に変換して出力する信号再生手段とを有し、第1の加算

4

手段で減算する推定背景雑音と第2の加算手段で減算する再更新推定背景雑音の大きさは、入力信号から背景雑音が除去されるように設定されることを特徴とする。

【0006】また、本発明は上述の課題を解決するために、入力信号として、背景雑音が入力され次いで該背景雑音と音声信号が混在する形で入力される背景雑音除去装置において、この装置は、入力信号を時間軸信号から順次フレーム単位で周波数成分に変換する信号変換手段と、入力信号に含まれる音声信号を検出する音声検出手段と、この音声検出手段で音声信号が検出されないノイズ期間は、信号変換手段で変換される周波数成分と1フレーム前に生成した推定背景雑音との平均をとることにより推定背景雑音を生成して保持する背景雑音推定手段と、音声検出手段で音声信号が検出される音声期間は、信号変換手段で変換される周波数成分から背景雑音推定手段で保持される推定背景雑音を減算する第1の加算手段と、この第1の加算手段の減算で得られる周波数成分を信号とし、背景雑音推定手段で保持される推定背景雑音を雑音として、全周波数帯域のS/Nと全周波数帯域を複数の分割した小領域毎のS/Nとを計算するS/N計算手段と、小領域のS/Nと全周波数帯域のS/Nの差が所定値以下の各小領域について、第1の加算手段の減算で得られる周波数成分と背景雑音推定手段で保持される推定背景雑音とを所定の割合で含む雑音を生成し、この雑音に背景雑音推定手段で保持される推定背景雑音を加算することにより再更新推定背景雑音を生成する帯域別背景雑音再推定手段と、信号変換手段の変換で得られる周波数成分から再更新推定背景雑音を減算する第2の加算手段と、この第2の加算手段の減算で得られるフレーム毎の周波数成分を時間軸信号に変換して出力する信号再生手段とを有し、第2の加算手段で減算する再更新推定背景雑音の大きさは、入力信号から背景雑音が除去されるように設定されることを特徴とする。

【0007】この場合、音声検出手段は、信号変換手段の変換で得られる周波数成分に含まれる音声成分を検出することにより入力信号に含まれる音声信号を検出するようにするとよい。

【0008】更に、音声検出手段は、信号変換手段の変換で得られる周波数成分を複数の小領域に分割して該小領域毎に音声成分を検出し、音声成分が検出された小領域の数が所定値以上であるとき入力信号に音声信号が含まれると判定するようにするとよい。

【0009】また、音声検出手段は、信号変換手段の変換で得られる周波数成分を複数の小領域に分割して該小領域毎に音声成分を検出し、全小領域数に対する音声成分が検出された小領域の数の割合が所定値以上であるとき入力信号に音声信号が含まれると判定するようにするとよい。

【0010】

【発明の実施の形態】次に添付図面を参照して本発明に

よる背景雑音除去装置の実施例を詳細に説明する。

【0011】図1は本発明の第1の実施例の背景雑音除去装置を示すブロック図である。この第1の実施例は、背景雑音から急峻な変化成分を取り除いた推定背景雑音および背景雑音の急峻な変化分を含む再更新推定背景雑音を生成し、入力信号から推定背景雑音を除去し、更に、周波数帯域を複数に分割した小領域のうちS/Nが特に小さい小領域に対しては再更新推定背景雑音を除去することにより、音質を損なうことなく背景雑音を精度よく除去するものである。

【0012】図1において、入力端子18には、始め背景雑音N1が入力され、次いで音声信号Sがその背景雑音N1に混在する形で入力されるものとする。例えば、携帯電話装置、自動車電話装置、音響機器等のマイクロホンから出力される信号がこの音声信号S、背景雑音N1に相当する。入力端子18に入力された入力信号（背景雑音N1、または音声信号Sと背景雑音N1）は、入力端子18に接続されるアナログ-デジタル変換器（A/D変換器）1に入力される。A/D変換器1は、入力端子18からの信号を所定の周波数、例えば、電話装置であれば8 KHzでサンプリングし、アナログ信号からデジタル信号に変換して出力側に接続される窓関数演算器2および時間軸音声検出器6へ出力する。

【0013】窓関数演算器2は、A/D変換器1からのデジタル信号を所定のサンプル数毎に1つのフレームにまとめ、その各フレームに対して窓関数演算を施して出力側に接続される高速フーリエ変換演算器（FFT演算器）3へ出力する。なお、窓関数演算には、フレーム間のデータ飛躍に基づく高周波成分の発生を防止するために、ハニング窓、ハミング窓等の窓関数が使用されることは周知のところである。FFT演算器3は、窓関数演算器2からの信号について1フレーム毎に高速フーリエ変換演算を施し、時間 t の関数として表わされる時間軸信号を周波数成分（スペクトル信号）に変換する。そして、そのスペクトル信号を、出力側に接続されるパワースペクトル計算部4および位相計算部8へ出力する。

【0014】ここで、入力端子18に背景雑音N1が入力されたときFFT演算器3から出力されるスペクトル信号を $N1(f, k)$ （ f は周波数、 k はフレーム番号）、音声信号Sのみが入力されたときFFT演算器3から出力されるスペクトル信号を $S(f, k)$ とすると、FFT演算器3から出力されるスペクトル信号 $X(f, k)$ は、入力端子18に背景雑音N1のみが入力されるノイズ期間では(1)式で表わすことができ、入力端子18に背景雑音N1と音声信号Sが混在する形で入力される音声期間では(2)式で表わすことができる。

【0015】

$$X(f, k) = N1(f, k) \quad (1)$$

$$X(f, k) = N1(f, k) + S(f, k) \quad (2)$$

パワースペクトル計算部4は、FFT演算器3からのスペ

クトル信号 $X(f, k)$ のパワースペクトル $X_{f, k}$ を、1フレーム毎に計算する。そして、計算したパワースペクトル $X_{f, k}$ を出力側に接続される加算器5および背景雑音推定部7へ出力する。位相計算部8は、FFT演算器3からのスペクトル信号 $X(f, k)$ の各周波数成分毎の同相成分および直交成分から各周波数成分の位相 $\Phi(f, k)$ を計算し、その位相 $\Phi(f, k)$ を出力側に接続される位相保持器9へ出力する。位相保持器9は、位相計算部8からの位相 $\Phi(f, k)$ を1フレームの期間保持する。

10 【0016】一方、時間軸音声検出器6は、A/D変換器1からのデジタル信号に音声が含まれるか否かを検出し、その検出結果を出力側に接続される背景雑音推定部7へ出力する。本実施例では、音声の検出を時間軸で行なうものであり、長時間、例えば500msにおける入力信号の平均パワーと、短時間、例えば50msにおける入力信号の平均パワーとの差を算出し、その差があらかじめ設定した閾値を越えたとき入力信号中に音声が含まれていると判定する。なお、音声の検出方法は、本実施例に限定されるものではなく、時間軸で音声を検出できる方法であればどのような方法であってもかまわない。

20 【0017】背景雑音推定部7は、スイッチ71、背景雑音更新部72、および背景雑音保持部73を含み、時間軸音声検出器6が音声を検出しないノイズ期間において、パワースペクトル計算部4からのパワースペクトル $X_{f, k}$ を用いて背景雑音を推定する。詳細には、スイッチ71は、時間軸音声検出器6からの判定結果に基づいてノイズ期間のみスイッチを閉じ、パワースペクトル計算部4の出力側を背景雑音更新部72の入力側に接続する。したがって、パワースペクトル計算部4からのパワースペクトル $X_{f, k}$ は、ノイズ期間のみ背景雑音更新部72に入力される。

30 【0018】背景雑音更新部72は、パワースペクトル計算部4からのパワースペクトル $X_{f, k}$ と、背景雑音保持部73に保持されている1フレーム前に推定した推定背景雑音 $N2_{f, k-1}$ とを用いて、(3)式により推定背景雑音 $N2_{f, k}$ を算出すると共に、背景雑音保持部73に保持されている1フレーム前の推定背景雑音 $N2_{f, k-1}$ を推定背景雑音 $N2_{f, k}$ に更新する。なお、(3)式の計算は、フレーム毎、周波数成分毎に実行される。

40 【0019】

$$N2_{f, k} = \alpha \cdot N2_{f, k-1} + (1 - \alpha) \cdot X_{f, k} \quad (3)$$

ここで、 $k=0$ のとき、 $N2_{f, 0}=0$ とする。また、 α は、ノイズ推定の速度を決定する係数であって、 $1 > \alpha > 0$ の範囲内に設定される。

【0020】ところで、背景雑音のパワースペクトルはその平均値の周辺に或る分散をもって時間的に変動するものである。したがって、背景雑音のパワースペクトルをそのまま背景雑音除去のために用いると、装置が各フレーム毎に背景雑音に過敏に応答し、その動作が不安定となる。そこで、本実施例では、(3)式を用いてパワー

7

スペクトル計算部4からのパワースペクトル $X_{f,k}$ と1フレーム前に推定した推定背景雑音 $N2_{f,k-1}$ との平均をとり、その平均したパワースペクトルを推定背景雑音として用いることにより、装置が各フレーム毎に背景雑音に敏感に応答することを防止し、装置の安定化を図っている。

【0021】このように、背景雑音推定部7は、ノイズ期間において、フレーム毎に(3)式により背景雑音を推定していく。そして、背景雑音保持部73で保持される推定背景雑音 $N2_{f,k-1}$ が、パワースペクトル計算部4から出力されるパワースペクトル $X_{f,k}$ に近づいて $N2_{f,k-1} \approx X_{f,k}$ となったとき、ノイズ期間では $X_{f,k} \approx N1_{f,k}$ ($N1_{f,k}$ は背景雑音のパワースペクトル)であるので、推定背景雑音 $N2_{f,k}$ を用いて背景雑音 $N1_{f,k}$ を除去することができることとなり、背景雑音推定部7の背景雑音の推定アルゴリズムは収束する。背景雑音更新部72は、推定アルゴリズム収束後の推定背景雑音 $N2_{f,k}$ を出力側に接続される加算器5およびパワー操作部10へ出力し、背景雑音保持部73は、その推定背景雑音 $N2_{f,k}$ を保持する。

【0022】ノイズ期間が終了し、次いで入力端子18に背景雑音 $N1$ と音声信号 S が混在する形で入力される音声期間に移ると、パワースペクトル計算部4からは背景雑音 $N1$ と音声信号 S が混在する信号に対応するパワースペクトル $X_{f,k}$ が出力される。加算器5は、このパワースペクトル $X_{f,k}$ と、背景雑音推定部7からの推定背景雑音 $N2_{f,k}$ とを用いて(4)式により減算を実行し、その結果 $E1_{f,k}$ を出力側に接続されるパワー操作部10および加算器13へ出力する。なお、(4)式において、 α_c は、推定背景雑音 $N2_{f,k}$ の減算量を加減するための係数であり、 $1 \geq \alpha_c > 0$ の範囲内で、使用環境等を考慮してあらかじめ

【0023】

$$E1_{f,k} = X_{f,k} - \alpha_c \cdot N2_{f,k} \quad (4)$$

先述したように音声期間では、FFT演算器3の出力 $X(f,k)$ は、 $X(f,k) = N1(f,k) + S(f,k)$ となるので、パワースペクトル計算部4から出力される音声のパワースペクトル

$$E1_{m,k} = \sum_{f=fs}^{fe} E1_{f,k}$$

$$N2_{m,k} = \sum_{f=fs}^{fe} N1_{f,k}$$

ここで、 m は、分割した小領域の周波数帯域の第 m 番目の周波数帯域(第 m チャンネル)を表わし、 fs は、第 m チャンネルにおける開始周波数を表わし、 fe は、第 m チャンネルにおける終了周波数を表わしている。

【0028】また、パワー計算部102は、加算器5からのパワースペクトル $E1_{f,k}$ 、背景雑音推定部7からのパワースペクトル $N2_{f,k}$ を用いて、(8)式、(9)式により、パワースペクトル $E1_{f,k}$ に対応する全周波数帯域に

8

ルを $S_{f,k}$ とすると、加算器5の出力 $E1_{f,k}$ は(5)式のようになる。

【0024】

$$E1_{f,k} \approx S_{f,k} \quad (5)$$

このように、加算器5からは、背景雑音を含む音声信号に対応するパワースペクトル $X_{f,k}$ から推定背景雑音 $N2_{f,k}$ を除去した音声信号に対応するパワースペクトル $S_{f,k}$ が出力されることになる。ただし、推定背景雑音 $N2_{f,k}$ は実際の背景雑音には完全に一致しないので、音声信号に対応するパワースペクトル $S_{f,k}$ のみを得ることは困難であり、 $E1_{f,k}$ には多少の背景雑音が含まれることとなる。なお、ノイズ期間では、FFT演算器3の出力 $X(f,k)$ は $X(f,k) = N(f,k)$ となるので、加算器5の出力 $E1_{f,k}$ は、 $E1_{f,k} \approx 0$ となる。

【0025】パワー操作部10は、帯域分割器101、パワー計算部102、対数計算部103、スイッチ104、およびパワー保持部105を含み、加算器5および背景雑音推定部7からのパワースペクトルを複数の小領域の周波数帯域(チャンネル)に分割し、そのチャンネル毎にパワーを計算するものである。詳細には、加算器5からのパワースペクトル $E1_{f,k}$ および背景雑音推定部7からのパワースペクトル $N2_{f,k}$ は、帯域分割器101およびパワー計算部102の双方にそれぞれ入力される。

【0026】帯域分割器101は、ノイズ期間においては背景雑音推定部7からのパワースペクトル $N2_{f,k}$ を、音声期間においては1フレーム毎に入力される加算器5からのパワースペクトル $E1_{f,k}$ をそれぞれ所定の数のチャンネルに分割し、出力側に接続されるパワー計算部102および後述する帯域別背景雑音再推定部12へ出力する。パワー計算部102は、帯域分割器101からの各チャンネルのパワースペクトルを用いて、(6)式、(7)式により、パワースペクトル $E1_{f,k}$ に対応する各チャンネルのパワー和 $E1_{m,k}$ 、パワースペクトル $N2_{f,k}$ に対応する各チャンネルのパワー和 $N2_{m,k}$ を計算する。

【0027】

【数1】

$$(6)$$

$$(7)$$

におけるパワー総和 $E1_{ALL}$ 、パワースペクトル $N2_{f,k}$ に対応する全周波数帯域におけるパワー総和 $N2_{ALL}$ を計算する。

【0029】

$$E1_{ALL} = \sum E1_{f,k} \quad (8)$$

$$N2_{ALL} = \sum N2_{f,k} \quad (9)$$

パワー計算部102は、計算したパワー和 $E1_{m,k}$ 、 $N2_{m,k}$ 、パワー総和 $E1_{ALL}$ 、 $N2_{ALL}$ を出力側に接続される

対数計算部103へ出力する。対数計算部103は、(10)～(13)式により、パワー計算部102からの各パワー和、パワー総和について対数変換を行い、出力側に接続されるスイッチ104へ出力する。

【0030】

$$E1_{mlog} = \log(E1_{m,k}) \quad (10)$$

$$N2_{mlog} = \log(N2_{m,k}) \quad (11)$$

$$E1_{alog} = \log(E1_{ALL}) \quad (12)$$

$$N2_{alog} = \log(N2_{ALL}) \quad (13)$$

スイッチ104は、時間軸音声検出器6により制御され、ノイズ期間はスイッチを端子a側に倒して対数計算部103の出力側をパワー保持器105に接続する。これにより、ノイズ期間では対数計算部103の出力はパワー保持器105に入力され、パワー保持器105は、対数計算部103からの各チャンネルのパワー和 $N2_{mlog}$ と全周波数帯域のパワー総和 $N2_{alog}$ を保持すると共に、出力側に接続されるS/N計算部11へ出力する。スイッチ104は、音声期間はスイッチを端子b側に倒して対数計算部103の出力側をS/N計算部11に接続する。これにより、音声期間では対数計算部103からの各チャンネルのパワー和 $E1_{mlog}$ と全周波数帯域のパワー総和 $E1_{alog}$ がフレーム毎にS/N計算部11に入力される。

【0031】S/N計算部11は、全帯域S/N計算部111および帯域別S/N計算部112を含み、パワー操作部10からの各チャンネルのパワー和 $E1_{mlog}$ 、 $N2_{mlog}$ 、全周波数帯域のパワー総和 $E1_{alog}$ 、 $N2_{alog}$ に基づいてS/Nを計算するものである。詳細には、全帯域S/N計算部111は、音声期間にスイッチ104を介して対数計算部103から送出される全周波数帯域のパワー総和 $E1_{alog}$ と、パワー保持器105に保持されているノイズ期間中の全周波数帯域のパワー総和 $N2_{alog}$ とを用いて(14)式により全帯域の S/N_a を計算し、出力側に接続される帯域別背景雑音再推定部12へ出力する。

【0032】

$$S/N_a = E1_{alog} - N2_{alog} \quad (14)$$

一方、帯域別S/N計算部112は、音声期間にスイッチ104を介して対数計算部103から出力される各チャンネルのパワー和 $E1_{mlog}$ と、パワー保持器105に保持されているノイズ期間中の各チャンネルのパワー和 $N2_{mlog}$ とを用いて、(15)式により小領域別の S/N_m を計算し、出力側に接続される帯域別背景雑音再推定部12へ出力する。

【0033】

$$S/N_m = E1_{mlog} - N2_{mlog} \quad (15)$$

ここで、mは、分割した複数の小領域の周波数帯域の第m番目の周波数帯域（第mチャンネル）を表わしている。

【0034】なお、本実施例では、パワー計算部102の出力を対数計算部103で対数変換した後、S/N計算部11でS/Nを算出しているが、対数変換しないでS/Nを算出してもよい。この場合、対数計算部103が不要となるが

S/Nの計算がやや面倒になる。

【0035】帯域別背景雑音再推定部12は、帯域別S/N差判定部121および帯域別背景雑音再更新部122を含み、S/N計算部11からの全周波数帯域の S/N_a および各チャンネルの S/N_m に基づいてチャンネル毎のS/N差を判定し、S/Nの低いチャンネルについては推定背景雑音の再更新を行なうものである。詳細には、帯域別S/N差判定部121は、S/N計算部11の全帯域S/N計算部111からの全周波数帯域の S/N_a および帯域別S/N計算部112からの各チャンネルの S/N_m を用いて、(16)式、(17)式により、チャンネル毎に S/N_m と S/N_a の差を判定し、その判定結果を出力側に接続される帯域別背景雑音再更新部122へ出力する。

【0036】

$$S/N_m \geq S/N_a + \delta \quad (16)$$

$$S/N_m < S/N_a + \delta \quad (17)$$

ここで、 δ は、装置や使用環境等を考慮してあらかじめ定めていく定数である。図2は、全周波数帯域の S/N_a とチャンネルの S/N_m の一例を示す図である。ただし、 $\alpha = 0.5$ 、 $\alpha_c = 0.7$ 、 $\beta = 0.8$ 、 $\delta = -6$ (dB)、 $m = 5$ の場合であり、また、縦軸の単位はdBである。

【0037】帯域別背景雑音再更新部122は、パワー操作部10の帯域分割器101から出力される各チャンネルの推定背景雑音のうち、(17)式を満足するチャンネルの推定背景雑音 $N2_{mf,k}$ と、同じく帯域分割器101から出力される加算器5の出力がチャンネル毎に分割された出力のうち、(17)式を満足するチャンネルの出力 $E1_{mf,k}$ とを用いて、(18)式により推定背景雑音 $N2_{mf,k}$ の再更新を行って再更新推定背景雑音 $N3_{mf,k}$ を生成し、出力側に接続される加算器13へ出力する。なお、(16)式を満足するチャンネルについては、再更新は行わない。

【0038】

$$N3_{mf,k} = (1 - \beta) \cdot N2_{mf,k} + \beta \cdot E1_{mf,k} \quad (18)$$

ここで、mは、分割した複数の周波数帯域の第m番目の周波数帯域（第mチャンネル）を表わす。また、 β は、再更新の速度を表わす係数であり、加算器5の第mチャンネルの出力 $E1_{mf,k}$ を再更新推定背景雑音 $N3_{mf,k}$ に反映させるためのものである。

【0039】(18)式において、 β を大きくすると再更新推定背景雑音 $N3_{mf,k}$ に対する $E1_{mf,k}$ の影響が大きくなり、 β を小さくするとその影響が小さくなる。ここで、加算器5の第mチャンネルの出力 $E1_{mf,k}$ は、背景雑音の急峻な変化に応じて急峻に変化するものである。したがって、 β が大きければ再更新推定背景雑音 $N3_{mf,k}$ は、急峻に変化する背景雑音成分に速やかに追従するものとなり、逆に β が小さければ再更新推定背景雑音 $N3_{mf,k}$ は、急峻に変化する背景雑音成分に緩やかに追従するものとなる。よって、 β は、再更新の速度を表わす係数であるということができ、使用環境等に応じて決定される。

【0040】加算器13は、帯域別背景雑音再更新部122

により推定背景雑音の再更新が行われるチャンネルに対して、帯域分割器101からの加算器5の第mチャンネルの出力 $E1_{mf,k}$ および帯域別背景雑音再更新部122からの再更新推定背景雑音 $N3_{mf,k}$ を用いて(19)式により減算を行い、出力側に接続される平滑演算部14へ出力する。

【0041】

$$\begin{aligned} E2_{mf,k} &= E1_{mf,k} - \alpha_k \cdot N3_{mf,k} \\ &= X_{mf,k} - \alpha_c \cdot N2_{mf,k} - \alpha_k \cdot N3_{mf,k} \end{aligned} \quad (19)$$

ここで、 α_k は、 $1 \geq \alpha_k > 0$ の範囲内であらかじめ定めておく係数であり、再更新推定背景雑音 $N3_{mf,k}$ の減算量を加減する係数である。

【0042】また、加算器13は、推定背景雑音の再更新が行われないチャンネルに対しては、(20)式に示すように加算器5の出力 $E1_{mf,k}$ をそのまま出力側に接続される平滑演算部15へ出力する。

【0043】

$$E2_{mf,k} = E1_{mf,k} \quad (20)$$

図3は、帯域別背景雑音再更新部の動作を説明するための図である。図3において、縦軸は各成分のパワーレベルを表わし、横軸はチャンネル(周波数)を表わしている。また、破線aは、背景雑音更新部72により推定された推定背景雑音 $N2_{f,k}$ のスペクトル、実線bは音声信号と背景雑音が混在したスペクトル $X_{f,k}$ 、点線cは背景雑音のスペクトル、点線dは帯域別背景雑音再更新部122で再更新された再更新推定背景雑音 $N3_{mf,k}$ のスペクトルである。また、斜線を施したチャンネル(4、6、9チャンネル)は、帯域別背景雑音再更新部122において推定背景雑音の再更新が行われたチャンネルである。

【0044】破線aの推定背景雑音 $N2_{f,k}$ は、(3)式で示されるようにフレームに対して平滑化された雑音であり、フレーム毎に急峻に変化する背景雑音成分に迅速に対応するものではない。したがって、加算器5において、点線bの音声信号と背景雑音が混在したスペクトルから破線aの推定背景雑音 $N2_{f,k}$ を除去しても、急峻に変化する背景雑音成分は残ってしまう。

【0045】そこで、(17)式に当てはまるS/Nの低いチャンネル、言い換えれば他のチャンネルに比較して音声成分が少ないチャンネル(図3の4、6、9チャンネル)については、加算器5の出力におけるそのチャンネルに対応する成分 $E1_{mf,k}$ を用いて急峻に変化する背景雑音成分に対応する点線dで示す再更新推定背景雑音 $N3_{mf,k}$ を(18)式により生成する。そして、加算器13において、(19)式により加算器5の出力 $E1_{f,k}$ からその再更新推定背景雑音 $N3_{mf,k}$ を除去する。これにより、加算器13の出力における音声成分が少ないチャンネル(図3の4、6、9チャンネル)のS/Nは、大幅に改善されることとなる。

【0046】平滑演算部14は、加算器13からの信号に対して(19)、(20)式による処理を施し、その信号の周波数

特性を平滑化する。この周波数特性の平滑化により各フレームの非連続性が減少するので、自然な音感を得ることができる。平滑演算部14は、周波数特性を平滑化した出力 $E3_{f,k}$ を出力側に接続される逆高速フーリエ変換演算器(逆FFT演算器)15へ出力する。

【0047】

$$E3_{f,k} = (1 - \gamma) \cdot E2_{f,k-1} + \gamma \cdot E2_{f,k} \quad (21)$$

逆FFT演算器15は、1フレーム毎に位相保持器9からの位相 $\Phi(f,k)$ を用いて平滑演算部14からの信号 $E3_{f,k}$ に対して逆高速フーリエ変換を施し、周波数成分で表わされる信号 $E3_{f,k}$ を時間tの関数として表わされる時間軸信号 $E4(k)$ に変換する。そして、出力側に接続される窓関数オーバーラップ処理部16へ出力する。なお、この周波数成分で表わされる信号を逆FFT演算により時間軸信号に変換する方法は既に公知である。

【0048】窓関数オーバーラップ処理部16は、逆FFT演算器15からの時間軸信号 $E4(k)$ を、フレーム間で不連続にならないようにあらかじめ定めた比率でオーバーラップさせることにより整合をとり、出力側に接続されるディジタル-アナログ変換器(D/A変換器)17へ出力する。D/A変換器17は、窓関数オーバーラップ処理部16からの信号をディジタル信号からアナログ信号に変換し、出力側に接続される出力端子19へ出力する。このようにして、入力端子18に入力された音声信号は、背景雑音が除去されて出力端子17から出力される。

【0049】以上説明したように第1の実施例によれば、入力信号を複数の小領域の周波数帯域(チャンネル)に分割して背景雑音の減算処理を行っているため、入力信号の特性が未知であっても処理内容に変更を加える必要がなく、また、1つの処理方法により低周波数域、高周波数域の分け隔てなく背景雑音の除去処理を行うので、聴感上の違和感が発生しない音質の優れた背景雑音除去装置を実現することができる。

【0050】また、第1の実施例によれば、入力信号を複数のチャンネルに分割し、各チャンネル毎に背景雑音の平均的な特性を示す推定背景雑音を用いて雑音減算処理を行い、他のチャンネルに比較してS/Nが小さいチャンネル、つまり音声成分の少ないチャンネルに対しては背景雑音の急峻な成分に対応する再更新推定背景雑音を用いて更に雑音減算処理を行っているため、音声成分が低周波数域に集中している場合であっても音声成分が少ないチャンネルで背景雑音を精度よく除去することができる。

【0051】図4は本発明の第2の実施例の背景雑音除去装置を示すブロック図である。この第2の実施例は、図1に示す第1の実施例の背景雑音除去装置における時間軸音声検出器6を周波数軸音声検出器20に変更し、パワースペクトル計算部4の出力側とパワー操作部10の帯域分割器101の出力側とを周波数軸音声検出器20の入力側にそれぞれ接続したものである。

【0052】図4において、周波数軸音声検出器20は、パワースペクトル計算部4の出力 $X_{f,k}$ と帯域分割器101の出力 $N_{2mf,k}$ を用いて、帯域分割器101において分割した小領域の周波数帯域（チャンネル）毎に(22)式により出力 $N_{4c,k}$ を計算する。

【0053】

$$N_{4cf,k} = \varepsilon \cdot N_{2cf,k} + (1 - \varepsilon) \cdot X_{cf,k} \quad (22)$$

ただし、 $N_{4cf,0} = 0$ とする。ここで、 c は、分割した複数の周波数帯域の第 c 番目の周波数帯域（第 c チャンネル）を表わし、 ε は、 $1 > \varepsilon > 0$ の範囲内であらかじめ

10

決めておくノイズ平均の速度を決定する定数である。

【0054】次いで、周波数軸音声検出器20は、(23)式により、(22)式により計算した出力 $N_{4cf,k}$ にあらかじめ設定した閾値 κ を加算する。そして、その加算結果 $J_{cf,k}$ が(24)式を満足する場合には、第 c チャンネルに音声成分が含まれるものと判定し、その加算結果 $J_{cf,k}$ が(25)式を満足する場合には、第 c チャンネルに音声成分が含まれないものと判定する。

【0055】

$$J_{cf,k} = \kappa + N_{4cf,k} \quad (23)$$

$$X_{cf,k} > J_{cf,k} \quad (24)$$

$$X_{cf,k} \leq J_{cf,k} \quad (25)$$

次いで、周波数軸音声検出器20は、音声成分が含まれていると判定したチャンネルの数 V があらかじめ定めたチャンネル数 M に対して(26)式を満足する場合には、第 k 番目のフレームは音声信号を含むフレームである判定し、(27)式を満足する場合には、第 k 番目のフレームは音声信号を含まないフレームである判定する。なお、チャンネル数 M は、特定の値に限定されるものではない。

【0056】

$$\text{音声成分有りのチャンネル数 } V \geq M \quad (26)$$

$$\text{音声成分有りのチャンネル数 } V < M \quad (27)$$

図5は、上述の周波数軸音声検出器20の動作を説明するための図である。図5において、縦軸はパワーレベル、横軸はチャンネル番号（周波数）である。また、実線 a は、パワースペクトル計算部4から出力される第 k 番目のフレームのパワースペクトル出力 $X_{f,k}$ 、実線 b は、背景雑音推定部7により推定された推定背景雑音のスペクトル $N_{2f,k}$ 、そして点線 c は、(23)式により計算された加算結果 $J_{f,k}$ を表わす。

【0057】この図5は、第2、第4、第6チャンネルについては(24)式が成立して音声成分が含まれるものと判定され、残る第1、第3、第5、第7～第10チャンネルについては(25)式が成立して音声成分が含まれないものと判定される場合の例である。この場合、音声成分が含まれるものと判定されるチャンネル数 V は3チャンネルであるから、例えば、(26)式、(27)式におけるチャンネル数 M を3とすれば、この第 K 番目のフレームは音声信号を含むものと判定される。なお、本実施例では、音声成分が含まれるチャンネル数 V によりフレームに音声

信号が含まれているか否かを判定したが、全チャンネル数に対する音声成分が含まれるチャンネル数 V の割合によりフレームに音声信号が含まれているか否かを判定してもよい。

【0058】周波数軸音声検出器20による判定結果は、第1の実施例の場合と同様に背景雑音推定部7のスイッチ71およびパワー操作部10のスイッチ104に送出され、各スイッチを制御する。なお、周波数軸音声検出器20以外の各部は、図1に示す第1の実施例の対応する各部と構成、動作が同じである。

【0059】以上説明したように第2の実施例によれば、音声信号の検出を周波数軸で行っているの、音声の周波数成分に重点を置いて検出し得る閾値の設定が可能である。また、フレーム単位で音声信号を検出するので多くのフレーム内のデータを一括して判定材料とすることになり、確実に音声信号を検出することができる。したがって、背景雑音を精度よく除去できる、音質の優れた背景雑音除去装置を提供できる。また、第2の実施例によれば、当然に第1の実施例と同様の効果を得ることができる。

20

【0060】図6は本発明の第3の実施例の背景雑音除去装置を示すブロック図である。図1に示す第1の実施例の背景雑音除去装置においては、音声期間において、パワースペクトル計算部4から出力されるパワースペクトル $X_{f,k}$ から、背景雑音推定部7で生成した推定背景雑音 $N_{2f,k}$ を加算器5で減算し、次いで、帯域別背景雑音再推定部12で生成した再更新推定背景雑音 $N_{3mf,k}$ を加算器13で減算することにより背景雑音の除去を行っている。

30

【0061】これに対して第3の実施例では、パワースペクトル計算部4から出力されるパワースペクトル $X_{f,k}$ から、推定背景雑音 $N_{2f,k}$ と再更新推定背景雑音 $N_{3mf,k}$ とをまとめて加算器13で減算することとしたものである。そのため、図6では、パワースペクトル計算部4の出力側を加算器5を介さずに直接加算器13の入力側に接続している。そして、帯域別背景雑音再推定部12は、再更新推定背景雑音 $N_{3mf,k}$ に一定量の推定背景雑音 $N_{2f,k}$ を加算した推定背景雑音 $(N_{3mf,k} + \alpha_c \cdot N_{2f,k})$ を生成し、出力側に接続される加算器13へ出力するものである。この第3の実施例においても、第1の実施例と同様の効果を得ることができる。

40

【0062】図7は本発明の第4の実施例の背景雑音除去装置を示すブロック図である。図4に示す第2の実施例の背景雑音除去装置においては、音声期間において、パワースペクトル計算部4から出力されるパワースペクトル $X_{f,k}$ から、背景雑音推定部7で生成した推定背景雑音 $N_{2f,k}$ を加算器5で減算し、次いで、帯域別背景雑音再推定部12で生成した再更新推定背景雑音 $N_{3mf,k}$ を加算器13で減算することにより背景雑音の除去を行っている。

50

【0063】これに対して第4の実施例では、パワースペクトル計算部4から出力されるパワースペクトル $X_{f,k}$ から、推定背景雑音 $N_{2f,k}$ と再更新推定背景雑音 $N_{3mf,k}$ とをまとめて加算器13で減算することとしたものである。そのため、図7では、パワースペクトル計算部4の出力側を加算器5を介さずに直接加算器13の入力側に接続している。そして、帯域別背景雑音再推定部12は、再更新推定背景雑音 $N_{3mf,k}$ に一定量の推定背景雑音 $N_{2f,k}$ を加算した推定背景雑音($N_{3mf,k} + \alpha_c \cdot N_{2f,k}$)を生成し、出力側に接続される加算器13へ出力するものである。この第4の実施例においても、第2の実施例と同様の効果を得ることができる。

【0064】なお、本発明は、携帯電話装置、自動車電話装置、音響装置等だけではなく、音声認識装置、TV会議装置、無線通信機等に広く適用することができる。

【0065】

【発明の効果】このように本発明によれば、入力信号の周波数帯域を複数の小領域に分割し、小領域毎に背景雑音の減算処理を行っているので、入力信号の特性が未知であっても処理内容に変更を加える必要がなく、また、1つの処理方法により低周波数域、高周波数域の分け隔てなく背景雑音の除去処理を行うので聴感上の違和感が発生しないという効果が得られる。

【0066】また、本発明によれば、小領域毎に、背景雑音の平均的な特性を示す推定背景雑音を用いて雑音減算処理を行い、更に、他のチャンネルに比較してS/Nが小さいチャンネルに対しては背景雑音の急峻的な変化に対応する再更新推定背景雑音を用いて雑音減算処理を行っているので、S/Nが小さいチャンネルにおける背景雑音を精度よく除去することができ、音質の優れた背景雑音除去装置を実現することができる。

【図面の簡単な説明】

【図1】本発明の第1の実施例を示すブロック図である。

【図2】チャンネルのS/Nと全周波数帯域のS/Nの一例を示す図である。

【図3】帯域別背景雑音再更新部の動作例を示す図である。

【図4】本発明の第2の実施例を示すブロック図である。

【図5】周波数軸音声検出器の動作例を示す図である。

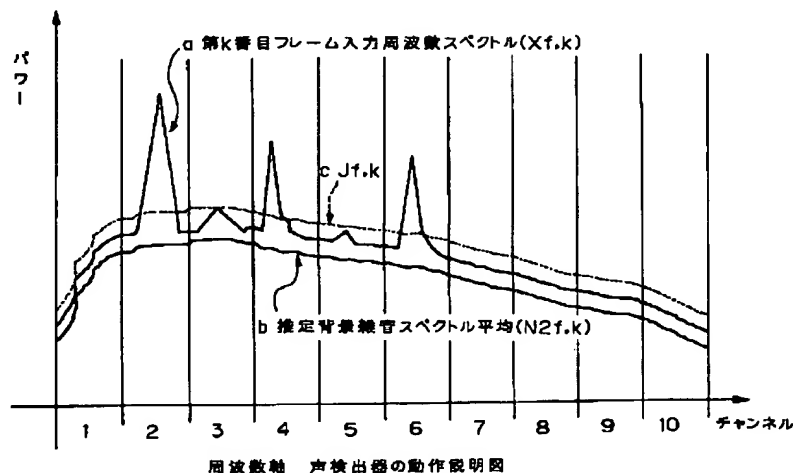
【図6】本発明の第3の実施例を示すブロック図である。

【図7】本発明の第4の実施例を示すブロック図である。

【符号の説明】

- 1 A/D変換器
- 2 窓関数演算器
- 3 FFT演算器
- 4 パワースペクトル計算部
- 5、13 加算器
- 6 時間軸音声検出器
- 7 背景雑音推定部
- 8 位相計算部
- 9 位相保持器
- 10 パワー操作部
- 11 S/N計算部
- 12 帯域別背景雑音再推定部
- 14 平滑演算部
- 15 逆FFT演算器
- 16 窓関数オーバーラップ処理部
- 17 D/A変換器
- 20 周波数軸音声検出器

【図5】



[illegible]

Figure 1 is a line graph showing the relationship between S/N (dB) and time. The y-axis is labeled S/N (dB) and ranges from -2 to 20. The x-axis is labeled 時間 (Time). The graph shows a sharp peak labeled S/N_m reaching approximately 18 dB, followed by a smaller peak labeled S/N_a reaching approximately 2 dB. The baseline is around -1 dB.

[illegible]

本発明の第3の実施例

